DOCUMENT RESUME

ED 353 305 TM 019 369

AUTHOR Stalford, Charles B.
TITLE Issues in Validating Effectiveness of School Improvement Programs.
PUB DATE Apr 92
NOTE 13p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, CA, April 20-24, 1992).
PUB TYPE Viewpoints (Opinion/Position Papers, Essays, etc.) (120) -- Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Documentation; *Educational Improvement; Effective Schools Research; Elementary Secondary Education; Federal Aid; *Federal Programs; Information Dissemination; Outcomes of Education; Program Development; Program Effectiveness; *Program Validation; *School Districts; Validity
IDENTIFIERS *National Diffusion Network Programs; Validation Verification and Testing Techniques

ABSTRACT

This paper argues that the National Diffusion Network (NDN) is a valuable tool for disseminating information about effective school improvement programs (SIPs). It considers why SIPs are not found more frequently in the NDN and what can be done to remedy this situation, from the perspective of the executive secretary of the NDN Program Effectiveness Panel (PEP). A multiplicity of activities occurs in SIPs. The goals of SIPs are quite broad, while those of NDN programs are content or subject-matter specific or relatively focused. Other contextual factors may impede movement toward gaining PEP approval (i.e., time and expense involved in collecting requisite evaluation data, and need to have a school or schools make the application). These traits of SIPs challenge their program sponsors to fully and clearly describe what the intervention is, document claims of significant outcomes that flow from the program, and attribute the observed outcomes to the program itself and rule out competing explanations. The Spencerport (New York) School District obtained PEP validation of a district-wide SIP by having a district-wide leadership team develop a district plan for school improvement based on effective schools research, establishing annual building plans, submitting district-wide test results, and collecting data on student outcomes and use of the effective schools processes over a sustained period. (RLC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED353305

Issues in Validating Effectiveness of School Improvement Programs

Paper presented at the Annual Meeting of the
American Educational Research Association
San Francisco, California, April 24, 1992

(Session 59-20: Results from Research-Based
School Improvement Programs)

Charles B. Stalford
Education Research Analyst
Office of Educational Research and Improvement
U.S. Department of Education

(These remarks will follow reactions I will present to the three progress reports presented in this symposium. My views do not necessarily reflect those of the Office of Educational Research and Improvement nor the U.S. Department of Education.)

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

CHARLES B. STALFORD

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM019369

BEST COPY AVAILABLE

Issues in Validating Effectiveness of School Improvement Programs

Introduction

I wish to focus attention on one issue as part of this symposium:

> What is the potential of the National Diffusion Network to serve as tool for dissemination of information about effective school improvement programs? To be more precise about my interest, I am particularly interested in the question "Why are school improvement programs not found more frequently in the NDN and what can be done about that?"

Background

I am particularly interested in these questions because I currently function as the executive secretary for NDN's Program Effectiveness Panel (PEP) within OERI. In my PEP capacity, I pre-review all submissions for panel approval (and occasionally sit as a member of individual PEP panels as well. I was also a member of PEP's predecessor, the Joint Dissemination Review Panel).

As you probably know, a program must be passed by PEP in order to be eligible for NDN Developer-Demonstrator (DD) funding, and other benefits of NDN membership. The procedures and evaluation criteria for PEP review of submissions are contained in the OERI publication Making the Case.[1] (About 60% of all submissions to PEP judged

according to Making the Case pass).

Factually, there are presently very few school improvement programs in NDN, and there are very few submissions currently from such programs to become members of NDN. In the judgement of staff in the contractor engaged to support NDN and PEP (the RMC Research Corporation), there are presently three such programs in NDN. By comparison, there are some 150 currently active projects listed in the 1992 edition of Educational Program that Work, the annual catalogue of NDN programs. Three school improvement programs out of 150--why aren't more in NDN?

Inasmuch as the prior presentations in this symposium have discussed school improvement programs, there is no point in my seeking to define them further here. For purposes of discussing PEP validation of school improvement programs, I wish to draw attention to two aspects of them. One is the multiplicity of activities that ordinarily occur in them. Thus, various sorts of teacher training, leadership development and other activities will be occurring more or less simultaneously. A multiplicity of activities may be found in many NDN projects as well; however, in general the scope of NDN projects tends to be more limited, for example, on instruction in a given subject area, or on processes affecting a certain set of grades and/or age levels.

The second potentially PEP-related aspect of school improvement programs is that, by definition, their goals tend to be quite broad, i.e., on school or district-wide improvement, not individual subjects or classrooms. In some respects the results may also be unpredictable. In part, that is because in individual school improvement programs, the choice of goals upon which to focus is left to the local participants themselves. In part, it is also apparently not always possible to predict what the combination of staff development, organizational and leadership activities implemented on a school or district-wide basis may ultimately yield.

This is in contrast to many NDN programs. Programs in NDN tend to be content or subject-matter specific or otherwise relatively focused, e.g., "Books and Beyond" (reading), "Comprehensive School Mathematics Program" (math), "Life Lab Science Program" (science). One will therefore often see in PEP submissions a program structure that says "If we do this intervention (in a particular area), we will these results (in the given area)".

These characteristics of school improvement programs pose the following three challenges for their program sponsors.

(1) A need to describe fully and clearly what the "intervention" is. This is important for PEP members to be able to judge whether the program could have brought about the claimed outcomes. It is also important to help PEP members

judge the replicability of the program. (Complexity of a program is not necessarily a bar to replication; lack of clarity about its essential components is.)

(2) A need to document claims of one or more significant outcomes that flow from the program. While one can not always predict the exact outcomes given the complexity of school improvement programs, it still should be possible, based on prior experience, to be able to say what outcomes can reasonably be expected.

(3) A need to be able to attribute the observed outcomes to the program itself, and rule out possibly competing explanations (e.g., that something else was going on concurrently).

There are other contextual factors which may impede movement toward gaining PEP approval. One of these, not unique to school improvement programs, is the time and expense involved in collecting requisite evaluation data. I can tell you from my vantage point of doing staff work for PEP that many potential applicants face serious obstacles in getting resources to conduct an evaluation adequate to support a PEP submission. (Those with long memories recall the days when Title III of the Elementary and Secondary Education Act, which supported demonstration school projects with attendant funding for program evaluation, enabled many projects to submit applications to PEP's predecessor, the

Joint Dissemination Review Panel in the 1970's.) Nonetheless, we currently continue to see submissions with high-quality evaluation designs and data.

There is another factor, perhaps more unique to school improvement programs, which I had frankly not considered until I had an informal conversation with one of today's presenters prior to this symposium. That is the need to have a school, or schools, make the application. The role of the laboratory in Onward to Excellence, for example, is to train the school to run its own program. The school therefore would ultimately be the submitter (to PEP).

There are variations on this theme, however, which possibly should be explored further. The Follow Through Program offers one approach. Follow Through has always had a structure in which a program developer (e.g., the High/Scope Foundation) acts as a sponsor for its model in numerous sites. The local school is therefore implementing the model with the continuing assistance of the program developer.

In the case of PEP, the model developer can be the submitter, present data from several schools implementing its model, and request validation of the whole model. The High/Scope Foundation this year in fact did exactly that, gaining PEP validation of both its preschool and K-3 curricula using data from several local sites in which the model(s) had been developed and evaluation data collected. There is no reason why one of the laboratories whose

programs have been discussed here, could not act in such a sponsoring capacity.

Again, the focus of my remarks is on the following questions:

> What is the potential of the National Diffusion Network to serve as tool for dissemination of information about effective school improvement programs? To be more precise about my interest, I am particularly interested in the question "Why are school improvement programs not found more frequently in the NDN and what can be done about that?"

I have discussed some issues more or less uniquely related to school improvement programs. I now wish to say that these issues are not insurmountable. Further, I believe that the general utility and acceptability--that, is the ultimate value-- of school improvement programs, including those discussed here today, would be enhanced if they were to become members of NDN.

And it can be done. The Spencerport, N.Y. school district has just succeeded in gaining PEP validation of a district-wide school improvement program. The program is closely based on the principles of the Edmonds and Lezotte.[2] In fact, the late Dr. Edmonds and Dr. Lezotte personally assisted the district with the implementation of this program.

Spencerport is a suburban district with some 3600 students. Its

school improvement program was developed in 1982-83, and has operated until the present time. It fully incorporates the effective schools tenets. A district-wide leadership team was formed. This team developed a district plan for school improvement based on the effective schools research. Building-level multi-role school improvement teams were then formed. Annual building plans concentrated on activities to bring about the correlates of effective schools in the buildings: clear school mission, high expectations for success, instructional leadership, a safe and orderly environment, frequent monitoring of student progress, opportunity to learn and student time on task, and home school relations. The program was said to be an ongoing process, not just an "event" or a "product", and one which would deal with "fundamental instructional issues" rather than any one of the correlates alone. Stated differently, in terms we are using today, this was a complex, multi-faceted intervention that posed challenges of the sort I discussed previously regarding PEP validation.

In its submission to PEP, Spencerport presented district-wide results on the Stanford Achievement Test (SAT) in math and science for seven years, beginning in 1982-83. The results showed a steady increase in students scoring above the 40th percentile, plus an increase in those scoring in stanines 7-9. In addition, the district showed improvements in student performance on the New York Regents examination, a curriculum-based assessment in specified subjects.

Other convincing data were presented to show that the correlates of effective schools were present in the schools in increasing numbers during this seven year period. The project was approved by PEP.

In its submission, Spencerport did several things that are worthy of consideration in our discussion. The first, and perhaps most noteworthy of them, is they collected data on both student outcomes and implementation of the effective schools processes over a sustained period of time.

Regarding the challenge posed by the multiplicity of events in an effective schools program, the district presented a clear picture of what they implemented, even thought complex, using the Edmonds-Lezotte effective schools principles as a referent. This was not, however, an easy part of the process. Prior to their validation, an earlier PEP panel had not approved the program, largely on the grounds that the intervention was not clearly described. In its ultimately successful re-submission, however, the district met this requirement handily.[3]

Spencerport took a high-risk approach to claiming student outcomes. They based their claim on district-wide standardized testing, plus results from the Regents exams. One might argue that, since this was a district-wide program, a district-wide measure such as standardized test results should be used for validation. On the other hand, the difficulties in using standardized tests as a measure in program evaluations are well known. This district did

so successfully.

To buttress its claim to significant student outcomes and rule out possible alternative hypotheses (always a PEP concern) data from five comparable districts were presented which did not show similar increases over the period of time studied. (In addition, the district, when asked by PEP, demonstrated that the increase did not start until after the program started, thereby ruling out another alternative hypothesis that a trend was already in place before the program's implementation.)

The second challenge I previously stated to PEP validation was that because the expected outcomes of school improvement program are broad, it can be difficult to measure them in an evaluation; also that, to some extent, that outcomes may be unpredictable. The breadth of the Spencerport data met that challenge: in addition to SAT results in reading and mathematics, the district presented Regents results in secondary math, earth science, physics and french. The results were sufficiently broad as well as focused on core academic subjects to be seen as significant on the face of it. They also suggest (but obviously do not guarantee) that an adopting district might expect to obtain the same breadth of results if it implemented the same breadth of programmatic features.

(PEP members sometimes wonder whether an applicant has formed its claims after the fact, i.e., by observing which, if any of the possible outcomes have been realized after the data have been

collected, rather than framing claims in advance. It is realized that everyone is dealing with real-world, somewhat unpredictable research in the field, not laboratory science. But a school-wide improvement program that presented evidence showing math scores improved, but reading scores did not, when there was no information to suggest that reading received any less attention in the program than did math, might raise such a question in the panelists' minds, or at least a question of why such a pattern existed. If there was information to show that the math component was effective in its own right, regardless of other activities being implemented, the PEP might approve the submission, but as a tool for math only. If there was no such clarifying information, the entire submission might be disapproved. In all such cases, the PEP is asking itself the question "what can an adopter of this program reasonably expect to achieve when implementing it in his or her own district?)

To suggest an answer to my original questions--(1) I think there is considerable value to education by having more school improvement programs in the NDN, particularly those in which considerable public funds have been invested. Validation procedures aside, the combination of program validation and dissemination activities supported by the NDN are a significant force for school improvement.

(2) Regarding my specific question about why more school improvement programs are not found more often in NDN, we have today discussed several reasons why that may be the case. But these

barriers can be overcome, as Spencerport illustrates. So I hope more programs will "take the plunge".

NOTES

1. U.S. Department of Education (OERI), Making the Case: Evidence of Program Effectiveness in Schools and Classrooms--Criteria and Guidelines for the U.S. Department of Education's Program Effectiveness Panel, Washington DC, November 1988.

2. For an expression of these principles, see: Edmonds, Ronald, "Effective Schools for the Urban Poor", Educational Leadership, Vol. 37, No.1, Oct. 1979, pp.15-23.

3. Under appropriate circumstances, applicants can resubmit programs to PEP for approval.